Pain Level Assessment through Facial Expressions and Physiological Signal Modalities Using Hybrid ML Techniques

Synopsis of the Research Proposal

Submitted for the Registration in the

Ph.D. Program (Part-Time) in

Computer Science and Engineering

By

Toshika Lata

(24RDCDIUJDD02022)

Under Supervision of

Prof. A. K. Saini

ICFAI Tech Jaipur

Table of Contents

Abstract..... 4

1. Introduction..... 5
1.1 Background of the Study..... 5
1.2 Existing Theories and Models of Pain..... 6
1.3 Theoretical Framework of the Present Study..... 7

2. Literature Review..... 9
2.1 Diagnosis of Pain Based on Facial Expressions..... 9
2.2 Pain Assessment Based on Physical Signs..... 9
2.3 Multimodal Pain Assessment.....10
2.4 Application of Hybrid Machine Learning and Deep Learning.....10
2.5 Research Gap 10

3. Present Study..... 12
3.1 Problem Statement..... 12
3.1.1 The Main Research Components..... 12
3.2 The study is important..... 12
3.3 Research Objectives..... 13
3.4 Research Questions..... 14

4. Proposed Research Methodology 15
4.1 Research Design..... 15
4.2 Research Strategy..... 15
4.3 Samples and Participants..... 17
4.3.1 Data for Training and Validation..... 17
4.3.2 Participants for Real-World Testing..... 16
4.4 Data Collection..... 16
4.4.1 Secondary Data..... 16
4.4.2 Data Collection Process..... 17
4.5 Ethical Considerations.....19
4.6 Analytical Analysis..... 19
4.6.1 Quantitative Analysis.....19
4.6.2 Advanced Analytical Tools..... 20
4.6.3 Model Performance Metrics.....20

5. Limitations of the study 21

6. Timeline and Work Plan..... 22

7. Expected Outcome..... 23
7.1 Main Expected Results..... 23
7.1.1 More Accurate Pain Detection 23
7.1.2 Better understanding of Pain 23
7.1.3 Automatic Pain Detection System 23
7.1.4 Helps for Doctors and Researchers 23
7.1.5 New Knowledge in Scientific 23

8. References..... 24

Abstract

Pain measurement is a significant part of healthcare because it determines the diagnostic and therapeutic choices. Widely used self-reported scales are subjective and greatly reliant on patient communication, and thus, they cannot be as reliable with nonverbal, pediatric, postoperative and cognitively impaired patients. As a result of this inconsistency, automated pain detection with the help of artificial intelligence and sensors have emerged.

Facial expression analysis has been widely researched on objective pain assessment [1], [3], [6], [8], [9]. Additional objective measurement of the intensity of pain is provided by heart rate, EMG, EEG, and physical measurements like GSR and electrodermal activity [4], [5], [10], [11]. Recent research indicates that the multiple fusion of the visual and physical systems is stronger compared to those whose sources are singular [3], [10], [15], [16].

Transformer-based, and attention-based fusion network models have demonstrated a promising future in deep learning and spatiotemporal architecture to enhance the ability in capturing complex temporal dynamics and detecting performance [14], [19], [21], [22], [24]. Nevertheless, the problems of generalizability, real-time performance, and clinical usefulness in other population groups are still observed [12], [13], [18].

This paper suggests a multivariate system of pain measurement based on a combination of physiological and facial cues (EEG, EMG, ECG, GSR) based on a network with temporal modeling. An attention-based fusion unit takes the place of combining facial and anatomical signals to classify pain optimally.

The BioVid and other datasets [2], [9] are subjected to strict preprocessing and cross-validation. The contributions are anticipated to be an in-depth knowledge of diagnostic accuracy, response to multiple pains and a scalable system that can be used in clinical settings in real time.

1. Introduction

It is a complex subject with both sensory and emotional aspects of pain; therefore, precise pain evaluation in clinical practice is very challenging. The conventional pain measures are mostly self-report, which are unreliable or unfeasible to unconscious, cognitively compromised, comatose, and non-verbal patients.

This is a significant drawback that has led to the designing of automated pain measurement systems based on artificial intelligence to aid the decision-making process in the healthcare domain [1], [12], [29]. The results of the research have indicated that pain is associated with recognisable facial expressions, head movements, physiological changes that give measurable signals in identifying objective pain [1], [5], [8].

Recent innovations in deep learning, multisensory models and large scale benchmark datasets such as BioVid Heat Pain database have created and further fueled progress in this sphere. Nonetheless, the availability of strong, precise, and generalised pain monitoring is still a significant concern.

1.1 Background of the Study

Results of research studies conducted in the previous years have indicated that pain brings about both visual and behavioural reactions in addition to physiological changes that can be measured. The major indicators of pain include facial expression, muscle activity, movement of the head, and movement of the eyes, which have been studied in a great number [1], [6], [8], [9], [18].

Simultaneously, autonomic nervous system responses to pain including heart rate, electrodermal activity, electromyography have been shown to be effective biomarkers of pain intensity [4], [5], [11], [23].

The contributions made to standardising evaluation of pain by BioVid, UNBC-McMaster and other multimodal datasets have made a significant contribution in the standardisation of evaluation, whereas a unimodal system still shows problems of variations in lighting, inhibition, emotions and the individual characteristics of person [2], [3], [9].

In order to counter such problems, recent studies have advocated a multi-component fusion of facial expression and biomarkers in an integrated manner to enhance their robustness and generalisability across different settings [10], [15], [27]. Despite the new developments in deep learning that have made it more efficient, a lack of diversity in the dataset and imbalance between classes, low interpretability, and constraints of real-time configuration are still unaddressed challenges [12], [14], [26].

The symptoms of mild and low-grade pains are especially hard to notice since they tend to have insidious manifestations as subtle facial muscle movements and debilitating physiological changes that are currently hard to detect in models [18], [27].

Moreover, their application cannot be scaled to remote or resource-constrained settings where no medical personnel are present, which is why automated and cost-efficient systems will be necessary to provide practical health assistance [13], [25].

These issues lead to the necessity of more credible and scalable pain assessment system that could be able to detect all degrees of intensity, including subtle pain, through the integration of facial expressions with several physiological cues.

Hence, this paper will apply Vision Transformer (ViT), ConvNext, and Temporal Deep Models to design a hybrid AI that can advance to develop a multimodal pain detection system based on attention-based fusion to enhance its classification capability, performance, and ability to work in real-time.

1.2 Theories and Models of pain that exist.

A lot of theories explain the nature of pain and ways in which it can be researched.

Facial Action Coding System (FACS): The system that defines the movement of the muscles of the face when a person is in pain [12], [13].

- **Biopsychosocial Model**: Demonstrates that physical, psychological, and social factors are relevant to pain [25].
- **Gate Control Theory**: According to which pain transmissions occur via a gate in the nervous system which can either enhance or reduce the signals until they reach the brain [26].
- **Neurophysiological Models**: demonstrate the impact of pain on physical processes, including heart rate, sweating, blood pressure and brain activity [14], [15].

The theories advocate the application of facial expressions and physical cue in precise determination of pain.

1.3 Theoretical Framework of the Current Research.

The research is founded upon the following:

- Biopsychosocial model [25]
- FACS Theory [12], [13]
- Neurophysiological model [13]

The study will examine facial expression and several physiological cues using an ML/DL model, such as ViT/ConvNext [22], [23], [24], [25] with LSTM/STA-LSTM architecture. Fusion methods of attention are used whereby facial and body cues are integrated to enhance the detection of pain. The Machine Learning (ML) models (SVM, Decision Trees) can be applied to small datasets [18].The Deep Learning (DL) models (CNN, Transformer, LSTM, ST-LSTM) are capable of automatically detecting features when applied to large datasets [19], [20]. The combination of ML and DL is a hybrid model that is more efficient, and attention-based fusion combines various features (spatial + temporal) to improve the classification accuracy [21], [22], [23].

2. Literature Review

2.1 The diagnosis of pain is based on the facial expression.

The use of facial expressions as meanings of pain has been discussed in a number of studies. The authors [12], [13] spoke about the way to analyse the facial, micro-expressions, and movements of muscles. Research suggested by the two studies showed improvement on the controlled conditions, but the research noted that real life conditioning can lower the accuracy by factors like lighting changes, camera angles and personal facial features.

There is some correlation with low expression or facial placement, preventing the assessment of pain and emphasizing the necessity to combine the analysis of faces with other physiological signs. In this part, the past research regarding visual and physical means of automatic pain evaluation is summarized. The summary has identified the progress of facial expression-based pain assessment systems, as an advanced, multidisciplinary, and hybrid deep learning framework, citing current limitations and future improvement opportunities.

2.2 Physical Sign-based pain assessment.

In [14], [15], [16], and [27], the physical signals that were studied and analyzed to detect pain included EEG, Ecg, EMG, GSR, PPG and respiration. It was proposed that these signals were used in the proposed model to give objective indicators of the intensity of pain and stress response. As an example, the brain activity correlating with the sensation of pain was measured using the EEG technique, whereas the GSR displayed the alterations in the skin conductance that are connected to the sympathetic nervous system.

Nonetheless, the current studies cited in [10], [11] observed that most of the studies relied on single-signal approaches, thus compromising the validity of pain detection. These models are also restricted in their generalizability by small and homogeneous datasets.

2.3 Multimodal Pain Assessment

More accurate and robust pain tests were enhanced by the traditional method of MRI and multimodal in the works of authors [17] [27]. The combination of facial expression and physical signs through smart algorithms has presented better outcomes. In the proposed study, some of the methods use two or more cues to enhance the ability to predict pain responses. The paper points out to the fact that the integration of heterogeneous data, handling of large data and use of these systems in clinical practice remains an issue.

2.4 Hybrid Machine Learning and Deep Learning Applications.

In research papers [22], [23], and [24], a novel hybrid methodology between the Deep Learning (DL) and the Machine Learning (ML) was introduced. This is achieved through attention-based fusion in order to bridge the limitations of single-signal models.

The approach combines Convolutional Neural Networks (CNNs) to extract local features of the facial image and temporal dynamics of body signals by the use of the sequential models, LSTM or ST-LSTM. The integrated ML and DL solution had an improved performance compared to the conventional solutions.

Research papers in [22], [23], [24] also revealed that these methods have been very few to apply on large and diverse populations and there is an evident gap in the research. The current study proposes addressing the problem of generalization and real-time implementation as detected in the previous researches by using the attention-based fusion to improve multi-signal pain detection and to remove the problem of generalization and real-time implementation.

2.5 Research Gaps

Literature reviews [17], [21], [22], [23], [24] indicate that the issues about pain research remain a challenge including:

Weaknesses:

Situations where single-model systems are used and the systems lack strong ability to handle crises.

- Unavailability of big, balanced, and diverse data.
- Weak research of hybrid deep learning models.
- Lack of real time and clinically usable infrastructure.

In an attempt to resolve these shortcomings, the current research is offering a hybrid ML and DL architecture-based generalized multivariate pain assessments framework. This method combines both facial and physiological responses via attention-based fusion, which tries to enhance precision, comprehension, and practical use.

3. Present Study

3.1 Problem Statement

Majority of the past research on automated pain measurement concentrated on a single form of data, i.e. facial expressions or body signals. It is due to this that the conclusions produced were less precise, particularly in real life scenarios. Moreover, the combination of facial and body cues is quite challenging as they are extremely different forms of information that are difficult to match in terms of time.

The research is evidently different. We require a universal system which is good at utilizing facial expressions and body cues. The hybrid deep learning model with attention based fusion will be used in the study as it performs well in the hospital and in general environments.

3.1.1 The principal elements of the research.

- Combine facial expressions and body gestures [12], [14], [15].
- The hybrid ML / DL model (VIT/ConvNext + LSTM / STA-LSTM) [22], [23], [24].
- Proper combination of vital signs through attention based fusion [22], [23], [24].
- Test the model on standard datasets e.g. BioVid Heat Pain [10], [11], [28], [29] and similar datasets

3.2 The study is important.

Failure to measure pain well may result in wrong treatments and damage to the patients [5], [6], [15], [16], [27]. Numerous former vision-based (eye or face-based) systems attempted to identify pain by using solely facial modes of expression. Nevertheless, they have a number of issues as they:

- Mostly used only one type of code(data).
- Were doing work on the collection of small data sets.
- The use of facial and bodily cues that are not used together properly [10], [11], [17], [21].

In this study, the issues will be solved by integrating both facial and body cues to come up with a more effective and concrete system [12], [13], [14], [15].

- A hybrid model (VIT / CONVINXT + LSTM / STA-LSTM) is taken.
- Machine learning (ML) is the most suitable option when it is applied to a small dataset.
- Deep Learning (DL) is more applicable with big and complicated data [22], [23], [24]].
- Attention-based fusion assists to distribute attention on key aspects of the model, hence enhancing accuracy and interpretation [22], [23], [24].

This research has three significances.

- **Clinical significance**: It will assist physicians in the correct timing to gauge pain [17], [27].
- **Scientific contribution**: It demonstrates the use of signals (facial and body) in a hybrid model of meditation [19], [20].
- **Practical application**: It can be applied to hospitals in different categories of patients [17], [27], [28], [29].

3.3 Research Objectives

- In order to augment the accuracy, reliability, and consistency of pain detection using facial expressions alongside body signals of EEG, ECG, EMG, GSR, PPG, and breathing [14], [15], [16], [17].
- To come up with a hybrid model (ViT /ConvNeXt + LSTM /STA-LSTM) that is efficient on small and large datasets [22], [23], [24].
- In order to apply the attention-based fusion to successfully incorporate the most significant features [22], [23], [24].

- In order to evaluate the system on the basis of common datasets like BioVid Heat Pain [10], [11].
- Try to come up with a real-time system that will be effective both within and without hospitals.

3.4 Research Questions

- Does the new model, which fuses facial expressions and body signals with the help of a hybrid ML/DL model, locate pain more accurately than more ancient solutions?
- Will this system be appropriate with various patients?
- Is the hybrid ML/DL faster and more accurate in real-time pain detection than single-signal or older AI systems?
- What is the benefit of attention-based fusion in mixing various signals and enhancing the comprehending of pain predictions?

4. The Proposed Research Method

This paper seeks to come up with a robotized system that will be able to measure pain levels concurrently using facial expressions and body language. This system will be real time in order to demonstrate the level of pain an individual is experiencing.

The system under development is based on a hybrid model of Machine Learning (ML) and Deep Learning (DL). It consists of models founded on transformers, attention fusion networks, and Explainable Artificial Intelligence (XAI) elements to ensure that the system is more accurate, interpretable, and clinically relevant in hospitals.

4.1 Research Design

This study has a mixed research design. This implies that this research entails quantitative (numerical) and qualitative (observational) elements.

Exploration Phase: The paper at this stage explores the association of facial micro-expressions, body signal variations, and pain stimulus [12], [14], [15].

Correlation Study: In this step, the relation between EEG, ECG, GSR, EMG, and PPG signals among other physiological parameters and self-reported pain intensities will be explored [5], [6].

Evaluation and Optimisation Phase: At this stage, different hybrid models are validated - ViT / ConvNeXt (facial features) and BiLSTM / STA-LSTM / TCN (body signals) [22], [23], [24].

4.2 Research Strategy

Quantitative Focus: The section of the work implies the application of multiple physiological indicators, including EEG, ECG, EMG, GSR, and PPG, and facial video to detect pain automatically [14], [15], [16].

Qualitative Support: The qualitative part aids in the comprehension of the emotional and behavioural side of pain using the words and responses of the participants [17], [27].

Structure of Study Phases

Data Collection Preprocessing Feature Extraction Model Training and Optimisation Validation Final Analysis [20].

Components of the Model

- Visual Transformer (ViT): Acquires and recognizes significant facial patterns and expressions.
- LSTM (Long-Short-Term Memory):
- BiLSTM (Bi-long short term memory):
- GRU (Gated Recurrent Unit):
- TCN (A Time-conscious Network):
- STA-LSTM (Scene Transformer with Augmented Long-term Memory):
- Record the responses of body signal and pain.

Fusion System

- **Attention-Based Fusion**: Weighs more weight on the most pertinent characteristics of each modality.
- **Transformer Fusion Network (TFN) and Multimodal Compact Bilinear Pooling (MCBP)**: Blend facial and body data characteristics.
- **Cross-Modal Self-Attention**: Combines geographical and time-related aspects to gain a more comprehensive insight into the pain.

Explainable AI (XAI) Tools: Grad-CAM and SHAP are applied to demonstrate the decision making by the model.

These tools enhance transparency and trust between the doctors and the researchers.

4.3 Samples and Participants

The training and validation data are data used to validate a trained model against the development data (training data).

The model is being trained and tested on the usual publicly-available datasets including: BioVid Heat Pain Database and other similar databases. Face expressions and physiology measures recorded in controlled conditions in the laboratory are present in these datasets [5], [6].

4.3.1 Testing Participants: The Real-World participants Adults of both genders able to safely receive mild pain [28], [29].

- **Sampling Method**: Non-probabilistic purposive sampling [28].
- **Cutoffs**: People with neurological, psychological, or facial disorders which might interfere with recordings [28], [29].
- **Sample Size**: The number of about 50-100 participants will be used, and it is statistically reliable [28], [29].

4.4 Data Collection

4.4.1 Secondary Data

The secondary data will come as a standard base of publicly accessible pain databases such as:

- BioVid Heat Pain Database and
- Similar Datasets

These datasets are multimodal data (facial expression and physiological measures) that were collected in controlled laboratory environments [10] -14].

4.4.2 Data Collection Process

Data Collection:

We have already received BioVid Heat Pain Dataset and trying to acquire more similar datasets from various dataset creators.

Preprocessing:

Eliminate motion and noise of devices with wavelet transformation and Butterworth filter. Normalize, synchronize all physiological signals.

Feature Extraction:

Isolate significant features in deep learning-based techniques, statistically based features and time-frequency domain features.

Model Training:

Train and optimize hybrid ML/DL models with the help of transfer learning, Bayesian optimization, and cross-validation.

Evaluation:

Determine model performance based on accuracy, precision, recall, F1-score, ROC, AUC, and Matthews Correlation Coefficient (MCC) [22] -[24].

4.5 Ethical Considerations

- That is, all of the information about participants will be secure and confidential.
- Informed consent will be gotten prior to any data collection.
- Research ethics will be in accordance to Ethical Committee, European Union (EU) and World Health Organization (WHO) rules [28].
- The privacy and security will be guaranteed by the adherence to the GDPR (General Data Protection Regulation) and the Indian Data Protection Laws.

4.6 Analytical Analysis

This paper will integrate quantitative data (signals and model outputs) and qualitative data (behavioural observations) to gain a better insight into human perception of pain.

4.6.1 Quantitative Analysis

Statistical Tests and Methods.

- Descriptive Statistics Mean, Standard Deviation (SD), Range.
- Normality Tests: Kolmogorov-Smirnov, Shapiro-Wilk Tests.

- There are two types, One-sample and Two-sample t-tests, which are known as parametric tests.
- All non-parametric tests: Wilcoxon Signed-Rank Test, Friedman Test.
- Post-hoc Tests: Bonferonni Correction, Tuckeys HSD.
- Effect Size: Cohen d, Eta squared (η 2), R.
- Correlation Analysis: Pearson/ Spearman correlation coefficient.

4.6.2 Advanced Analytical Tools

Type of Features: Rank and determine important features.
AHP (Analytic Hierarchy Process): Find out the relative significance of variables.
MANOVA (Multivariate Analysis of Variance): test the interaction of multiple variables.
Feature Significance: Consider features, significance, size and time frame.
Model Comparison Compare models: SVM, Random Forest, CNN, ViT, LSTM, STA-LSTM, and TCN.

4.6.3 Model Performance Metrics

- F1-Score
- ROC–AUC Curve
- Matthews Correlation Coefficient (MCC)
- Balanced Accuracy

5. Limitations of the Study

- In this study, there are no vocal or verbal expressions involving pain.
- The research is not based on chronic or long-term pain conditions, but only on acute (short-term) pain.

6. Timeline and Work Plan

Table 1 : Time Schedule of Research

Task	Details	Duration
Planning and Preparation	In this phase, the whole research work will be planned. The main objectives will be set. Tools like **Python, TensorFlow, and PyTorch** will be selected. Datasets such as **BioVid** and **X-ITE** will be chosen. With the help of the supervisor, a full research roadmap will be prepared.	1 Year
Data Collection and Literature Review	Review previous studies related to pain detection. Collect and clean data of facial expressions and body signals using standard datasets such as **BioVid**. Find main research gaps and select suitable **ML/DL models** for further work.	3 Months
Design of Methodology	Check the collected data and decide the final methods for **preprocessing** and **feature extraction**. Design the **hybrid model (ViT / ConvNeXt + LSTM / STA-LSTM)** with **attention-based fusion** for accurate pain prediction.	3 Months
Implementation	Build and test the designed hybrid model. Combine facial and physiological signals. Train and test the model using accuracy, precision, recall, and F1-score to check performance.	6 Months
Experimentation and Validation	Perform large-scale testing and **cross-validation**. Compare the hybrid model with other models like **SVM, CNN, LSTM, and TCN**. Study the results to understand how the attention-based model improves pain detection.	3 Months
Final Compilation and Thesis Submission	Present all results and findings to the supervisor. Make corrections as per feedback. Prepare and submit the final report and complete thesis for evaluation.	3 Months

7. Expected Outcome

This study will help to make a better and more correct way to find how much pain a person feels. It will use both **face expressions** and **body signals** to know the level of pain. By using **Machine Learning (ML)** and **Deep Learning (DL)**, the system will learn how the **face and body react together** when a person is in pain.

7.1 Main Expected Results

7.1.1 More Accurate Pain Detection:

The new system will find and measure pain more correctly than old methods, which only depend on what the person says or on one type of signal.

7.1.2 Better Understanding of Pain:

By using many signals like **EEG, ECG, EMG, GSR, PPG, and breathing**, the model will show a more clear and full picture of how the body reacts when a person feels pain.

7.1.3 Automatic Pain Detection System:

A computer system will be made that can automatically tell the pain level by using face expressions and body signals.

7.1.4 Help for Doctors and Researchers:

The results will help doctors and researchers to understand pain better. This will help them to make more correct diagnosis and better treatment plans.

7.1.5 New Knowledge in Science:

This study will add new information in the field of pain research. It will bring a new hybrid ML/DL system that can check pain in a complete and scientific way.

8. References

1. Philipp Werner, Ayoub Al-Hamadi, Robert Niese, Steffen Walter, Sascha Gruss, and Harald C. Traue, "Towards Pain Monitoring: Facial Expression, Head Pose, a New Database, an Automatic System and Remaining Challenges," *Proc. British Machine Vision Conference (BMVC)*, 2013.
2. Steffen Walter, Philipp Werner, Sascha Gruss, Harald C. Traue, Ayoub Al-Hamadi, et al., "The BioVid Heat Pain Database: Data for the Advancement and Systematic Validation of an Automated Pain Recognition System," *Proc. IEEE Int'l Conf. on Cybernetics*, 2013.
3. Philipp Werner, Ayoub Al-Hamadi, Robert Niese, Steffen Walter, Sascha Gruss, and Harald C. Traue, "Automatic Pain Recognition from Video and Biomedical Signals," *Proc. Int'l Conf. on Pattern Recognition (ICPR)*, 2014.
4. Philipp Werner, Ayoub Al-Hamadi, Steffen Walter, Sascha Gruss, and Harald C. Traue, "Automatic Heart Rate Estimation from Painful Faces," *Proc. Int'l Conf. on Image Processing (ICIP)*, 2014.
5. S. Walter, S. Gruss, K. Limbrecht-Ecklundt, H. C. Traue, P. Werner, A. Al-Hamadi, et al., "Automatic Pain Quantification Using Autonomic Parameters," *Psychology and Neuroscience*, 2014.
6. P. Werner, A. Al-Hamadi, K. Limbrecht-Ecklundt, S. Walter, S. Gruss, and H. C. Traue, "Automatic Pain Assessment with Facial Activity Descriptors," *IEEE Transactions on Affective Computing*, 8(3), 2017.
7. P. Werner, A. Al-Hamadi, and S. Walter, "Analysis of Facial Expressiveness During Experimentally Induced Heat Pain," *Int'l Conf. on Affective Computing and Intelligent Interaction Workshops (ACIIW)*, 2017.
8. P. Werner, A. Al-Hamadi, K. Limbrecht-Ecklundt, S. Walter, and H. C. Traue, "Head Movements and Postures as Pain Behavior," *PLOS ONE*, 13(2), 2018.
9. P. Lucey, J. Cohn, K. Prkachin, and P. Solomon, "Painful Data: The UNBC-McMaster Shoulder Pain Expression Archive Database," *Proc. IEEE Int. Conf. on Automatic Face & Gesture Recognition*, Santa Barbara, CA, USA, 2021, pp. 57–64. DOI: 10.1109/AFGR.2011.5771480.
10. M. Benavent-Lledo et al., "A Comprehensive Study on Pain Assessment from Multimodal Sensor Data," *Sensors (MDPI)*, 2023. DOI: 10.3390/s23249675.
11. R. Fernandez Rojas et al., "Multimodal Physiological Sensing for the Assessment of Acute Pain," *Frontiers in Pain Research*, 2023. DOI: 10.3389/fpain.2023.1150264.
12. M. Cascella et al., "Artificial Intelligence for Automatic Pain Assessment: Research Methods and Perspectives," *Pain Research and Management (Hindawi)*, 2023. DOI: 10.1155/2023/6018736.

13. M. Bellal et al., “Calibration Trial of an Innovative Medical Device (NEVVA©) for the Evaluation of Pain in Non-Communicating Patients in the Intensive Care Unit,” *Frontiers in Medicine*, 2024. DOI: 10.3389/fmed.2024.1309720.
14. S. Friedrich et al., “Faces of Experimental Pain: Transferability of Deep Learned Heat Pain Features to Electrical Pain,” *arXiv preprint*, arXiv:2406.11808v1, 2024.
15. Z. Hammal et al., “Automated Assessment of Pain (AAP) and Multimodal Sensing Grand Challenge for Next-Gen Pain Assessment (AI4Pain),” *Proc. Int. Conf. Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW)*, Sep. 2024. DOI: 10.1109/ACIIW63320.2024.00005.
16. R. Gamarra, J. Guerrero-Osuna, and W. Villegas-Ch, “Multimodal AI Techniques for Pain Detection: Integrating Facial Gesture and Paralanguage Analysis,” *Frontiers in Computer Science*, Jul. 2024. DOI: 10.3389/fcomp.2024.1420541.
17. E.-H. Jang et al., “Comparing Multimodal Physiological Responses to Social and Physical Pain in Healthy Participants,” *Frontiers in Public Health*, 2024. DOI: 10.3389/fpubh.2024.1387056.
18. N. B. Aoun, “A Review of Automatic Pain Assessment from Facial Information Using Machine Learning,” *Technologies (MDPI)*, 2024. DOI: 10.3390/technologies12060092.
19. M.-D. Nguyen et al., “Transformer with Leveraged Masked Autoencoder for Video-Based Pain Assessment,” *arXiv preprint*, arXiv:2409.05088v3, 2024.
20. A. Ridouan, A. Bohi, and Y. Mourchid, “Improving Pain Classification Using Spatio-Temporal Deep Learning Approaches with Facial Expressions,” *GitHub repository*, 2024–2025.
21. M. Safarzadeh et al., “Advancing Pain Assessment Through Multimodal Facial Sequential Image Analysis: A Transformer-Based Approach,” *SSRN Preprint*, 2024–2025. Available: https://ssrn.com/abstract=5201152.
22. S. Duan et al., “QAFE-Net: Quality Assessment of Facial Expressions with Landmark Heatmaps,” *Proc. IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, Jan. 2024. DOI: 10.48550/arXiv.2312.00856.
23. E. Pavlidou and M. Tsiknakis, “Multimodal Pain Assessment Based on Physiological Biosignals: The Impact of Demographic Factors on Perception and Sensitivity,” *Proc. 11th Int. Conf. ICT4AWE*, 2025. DOI: 10.5220/0013426800003938.
24. C. W. Tan et al., “Automated Pain Detection Using Facial Expression in Adult Patients with a Customized Spatial Temporal Attention LSTM Network,” *Scientific Reports*, Springer Nature, 2025. DOI: 10.1038/s41598-025-97885-5.
25. A. Subramanian et al., “Multimodal Pain Recognition in Postoperative Patients: Machine Learning Approach,” *JMIR Formative Research*, 2025. DOI: 10.2196/67969.
26. R. Fang et al., “Survey on Pain Detection Using Machine Learning Models: Narrative Review,” *JMIR AI*, 2025. DOI: 10.2196/53026.

27. S. Gkikas et al., “Multimodal Automatic Assessment of Acute Pain Through Facial Videos and Heart Rate Signals Utilizing Transformer-Based Architectures,” *Frontiers in Pain Research,* 2024. DOI: 10.3389/fpain.2024.1372814.

28. S. Gkikas, R. Fernandez Rojas, and M. Tsiknakis, “PainFormer: A Vision Foundation Model for Automatic Pain Assessment,” *arXiv preprint,* arXiv:2505.01571v3, 2025. Available: https://github.com/GkikasStefanos/PainFormer.

29. M. Cascella et al., “Artificial Intelligence for Automatic Pain Assessment: Research Methods and Perspectives,” *Pain Research and Management,* 2023. DOI: 10.1155/2023/6018736.